

## Data Swagger: A Systemic Approach to Train, Motivate and Engage Data Savvy Employees

Kaveh Abhari  
DiLab, San Diego State University  
[kabhari@sdsu.edu](mailto:kabhari@sdsu.edu)

Robert Brodskiy  
DiLab, San Diego State University  
[rbrodskiy0155@sdsu.edu](mailto:rbrodskiy0155@sdsu.edu)

Darshan Davis  
DiLab, San Diego State University  
[ddavis7833@sdsu.edu](mailto:ddavis7833@sdsu.edu)

Harris Ness  
DiLab, San Diego State University  
[hness7310@sdsu.edu](mailto:hness7310@sdsu.edu)

Janmae Pagador  
DiLab, San Diego State University  
[jpagador@sdsu.edu](mailto:jpagador@sdsu.edu)

Mikay Parsons  
DiLab, San Diego State University  
[miparsons@sdsu.edu](mailto:miparsons@sdsu.edu)

### Abstract

*The relevance of data literacy has increased substantially over the past three decades. When trained well, data-literate employees at all levels can make data-driven decisions, improving the overall performance of their organization. Utilizing Transformative Learning Theory (TLT) and Experiential Learning Theory (ELT), this paper proposes a systematic data education framework for increasing data literacy across organizations. Focusing on the needs and experiences of non-expert end-users, this model proposes the following four learning strategies in data literacy training design: experiential data training, critical incident reflection, rational open discourse, and autonomous experimentation. To inform this model and further investigate barriers to data literacy in organizations, interviews were conducted with individuals from two different data analytics units in the U.S. Department of Defense. This research provides key insight and practical suggestions for developing and improving data literacy training programs.*

### 1 Introduction

The emergence of data in our digitalized society has led to better business decisions, more organized work structures, and a more consistent flow of information [1]. These improvements are key to the competitive success of many organizations seeking to stay ahead in the new digital world [2]. Although this fact is widely recognized, useful data are stuck inside information silos and thus cannot be used by data-illiterate employees [3]. The unprecedented amount of data collected by organizations in the status quo necessitates employees develop data literacy. In addition, due to the vast data-related resources

available, employees must now take the initiative to learn the nuances of relevant programs and software. Although current data initiatives train employees on how to use data to make procedural decisions, they fall short in motivating and empowering individuals to self-regulate data usage. Without this motivation, few employees adopt the practice of making data-driven decisions. The present study argues that this is a consequence of traditional training models that solely utilize data use cases over teaching general data literacy and nuanced understanding. Addressing this limitation, the present study attempts to create an improved systematic education model for enhancing data literacy across organizations.

This study proposes a framework of data education to help organizations support end-users—those that use the information provided by data analysts and scientists—in understanding both the value and uses of data. To do so, interviews were conducted with members of two data analytics units within the U.S. Department of Defense. These interviews provided insight into the barriers to and functions of data literacy in their respective departments, corroborating the training framework proposed by the present study. In addition to insight provided by these interviews, the proposed data education framework was adapted from the Transformative Learning Theory (TLT) [4] and Experiential Learning Theory (ELT) [5], [6]. When combined, these theories offer a more desirable framework for data education. The TLT informs effective strategies to motivate and facilitate adult learning through promoting learning agency and adult reasoning. Relevant to the study at hand, the TLT can be contextualized to self-regulated learning environments like those faced by employees today [7]. While TLT emphasizes self-driven critical reasoning,

ELT argues the importance of learning from experience in a natural, multi-dimensional learning environment [5]. Previous research has shown that the TLT is favored by employees due to its presentation of information they perceive to be credible. However, ELT as a framework for teaching data literacy promotes learning through doing, allowing individuals to independently stay up to date with evolving technology. The proposed model incorporates components of both TLT and ELT to create one cohesive framework for increasing data literacy amongst employees.

This paper first presents research highlighting the value of data literacy, followed by an explanation of the main barriers to increasing data literacy faced by organizations in the status quo. The results of the case study are then presented, followed by the proposed model for data literacy education. The paper concludes with limitations and suggestions for future research.

## 2 Background

This study conceptualizes data literacy as the awareness of, curiosity for, and competencies related to the use of data in an organizational setting [8]. Data literacy is associated with different competencies, including the ability to find, understand, organize, and evaluate data collected over time. While data literacy is not solely captured by an individual's data-capabilities, familiarity and comfort with data are prerequisites to data-driven decision-making. Employees with data literacy competencies protect organizations against the weaponization of data and cultivation of misinformation, improving the overall function and success of their teams [9].

Because of the increased availability of data and the evolution of business intelligence tools, data skills are widely seen as integral to the future workplace [10]. As data become more prominent, employees are expected to become competent in areas like query development, visualization, and quality control. Being able to understand and interpret collected data is becoming a necessity in many fields because of the development of new technologies like dynamic visualization, self-service analytics, and non-code machine learning. New technological developments offer possibilities for most employees to make data-driven decisions not previously possible. As a result, data-literate employees can use data to gain an unmatched level of insight that supports better performance overall [11][12].

With improved training strategies, data literacy can increase organizational success. Therefore, as data science continues to rise in relevancy, employers

should prioritize supporting *all* employees in gaining the key competencies associated with data literacy.

### 2.1 Three Levels of Literacy

Data literacy is key to successful digital transformation and globalization. However, data competency demands are contextual to one's role in an organization. This study recognizes the differences between the definition of data literacy among two groups: *data analysts/scientists* and *end-users*. First, data analysts are responsible for processing data into valuable information for end-users. The results of these analyses must then be understood by end-users for them to make data-driven decisions with the highest possibility of success. For data analysts, data literacy involves competencies related to data science from gathering and organizing data to engaging in analyses like prediction, trend analysis, and recommendation development [13][14]. Knowing how to effectively utilize different programs and applications like *Python*, *SQL*, *R*, *Tableau* and *Microsoft PowerBI* is becoming a key component of data literacy for data analysts. Working with bigger and more complex data also requires familiarity with tools that allow for easier collection, processing, and interpretation. For example, technologies such as *Google BigQuery*, *Apache Spark*, and *AWS Redshift* allow analysts to access, organize, and analyze large amounts of data. More importantly, literacy in data analytics must include competencies related to efficiently and effectively communicating results insights from data analyses with the relevant parties.

The level of literacy in data science is what differentiates data analysts from data scientists in modern organizations. While data analysts play a significant role in helping businesses process structured business data and report on current trends, data scientists focus on complex data mining for advanced predictive and prescriptive analytics. Data scientists also specialize in designing data modeling processes that extract insight from unstructured data. Data literacy at this level involves competencies associated with behavioral and social sciences, industrial engineering, computer science, and advanced visualization and representation [8][15]. For example, data scientists are familiar with the level of sophistication involved in big data processing and automation frameworks [16]. In addition to industry knowledge, a strong educational background is usually required to develop the depth of knowledge necessary to be a data scientist.

At the end-user level, data literacy is associated with non-specialists who utilize extracted data to make informed decisions. Promoting data literacy amongst all employees enhances an organization's capacity to

understand the value of data and make data-driven decisions. The present study focuses on end-users, as end-users often do not have the data literacy necessary to utilize data or data tools to their full extent. Motivating, engaging, and training end-users is critical to the success of institutional data initiatives. Data literacy at the end-user level is also closely related to ICT (Information and Communication Technology) literacy, which is defined as using “digital technology, communication tools, and or networks to access, manage, integrate, evaluate, and create information in order to function in a knowledge society.” This definition mirrors the conceptualization of data literacy used in the present study [17].

The success of data initiatives highly depends on the end-user’s adoption of data tools and products developed by the data analysts and data scientists, as well as their involvement in their development, refinement, and maintenance. End-users may develop data literacy at three levels. The first level is basic data literacy, in which a mix of cognitive and technical skills are used to integrate data [8][15]. At this level, if successful, employees can perform common tasks such as data entry and disseminating data information relevant to their work. These are non-specialists who utilize extracted data to make informed decisions. End-users, who will be referenced throughout this research, commonly fall in this range. The second level is a medium level of data literacy, where employees can successfully manage and evaluate data. These employees can perform simple data analysis and understand the relevancy of certain data to their work. At this level, employees can make judgments about the “quality, relevance, usefulness, and efficiency” of data. The third level is advanced literacy in which employees can create knowledge based on technical tools in addition to industry knowledge. Advanced data literacy requires a strong educational background to develop the depth of knowledge required.

## **2.2 The Value of Data Literacy**

The role of data continues to increase in modern organizations, as employees are expected to make data-informed decisions rather than relying on intuition or waiting for instruction [18]. Hence, organizational data literacy drives value in a variety of ways. The value of organizational data literacy comes from the measurement of any financial or non-financial benefits that come from data-driven decisions—for example, new product decisions, customer service improvement, cost saving, and error reduction [19]. Organizational data literacy can also increase efficient and effective operations. As a result, organizations may save valuable resources and maximize their productivity with direct or indirect

benefits like greater profit and employee satisfaction, respectively. In addition, individual employees experience benefits when they can access and use organizational data such as trust, transparency, engagement, and self-efficacy.

In order to capitalize on the value of data, modern organizations look for those with the right set of qualifications, including data literacy. These qualifications are not limited to understanding the value of data or examining the quality of data sources. New data savvy employees are expected to have a general understanding of how data are generated, stored, and managed [20]. This level of data literacy allows end-users to make more informed judgments about the source, quality, and potential applications of datasets, as well as the challenges related to their variation, volume, and velocity. The first step toward having all employees participate in data-driven decision-making is to help end-users understand the potential value of existing datasets and realize those values using data products such as KPI dashboards and optimization tools. However, ‘the first step is always the hardest.’ Modern organizations face different obstacles to enhance data literacy organizational-wide. The next section explores these challenges from both operational and employee perspectives, followed by a discussion of the role of data literacy training in overcoming these challenges.

## **2.3 The Challenges with Data Literacy**

The above sections demonstrate the integral role data-literate end-users play in enhancing organizational operations. However, organizations face many challenges when attempting to increase data literacy amongst users less familiar with data in the first place. Part of managing data literacy at an operational level is developing data education plans responsive to the needs and skillsets of each unique end-user. Currently, managers cannot identify their teams’ needs, much less provide them with the necessary training to leverage data in day-to-day operations [1].

End-users face many difficulties in their quest to gain data literacy. First, making data presentable and consumable is key to data-driven decision making. As the need for data emerges rapidly, so do the tools associated with it. Therefore, employees will always need to learn new concepts to efficiently mine through data. For example, those who are not accustomed to seeing data in the form of a dynamic business dashboards may have difficulty assessing the relevancy of the data [8]. In addition, end-users must develop the skills to verify the quality of data. Both managers and employees face this difficulty as high quantities of data are collected. Hence, managers can

become highly dependent on their teams to evaluate data before any application. In terms of the applicability of data, a common problem faced by end-users is understanding which data are relevant to their operational goals. Even then, once the relevancy is determined, users may not understand if it is feasible to use [21]. This lack of understanding keeps managers and employees alike away from using data or tools.

Some end-users already have the knowledge base to use data and business intelligence tools. However, many still struggle to apply these deliverables to solve problems. These problems lead to lower adoption rates, which in turn decreases productivity and efficiency. As found in our study, even when these challenges are addressed at the operational level, employees often prefer to make experience-driven decisions data driven decisions.

The challenges faced by organizations attempting to increase data literacy amongst end-users are directly related to organizational culture. We argue that data literacy initiatives can benefit from organizational cultures that promote critical thinking and knowledge discovery with data. The lack of such culture is thus also a barrier to improving organizational data literacy. Thus, the present study proposes a training framework that has the potential to increase the value of learning within organizations by encouraging employees to continuously engage in self-directed data education.

## 2.4 Data Literacy Education

Creating a data literacy education plan prioritizes learning within an organization, and can thus motivate employees to develop the data literacy skills necessary to use data in decision-making. This requires that leaders of data initiatives assess employees' attitudes toward data applications, as well as their educational needs in terms of data access, data use cases, data tools, and data management. This affects how employees react to these educational programs and ultimately potential organization benefits. Therefore, creating educational programs with reasonable flexibility that can meet the needs of different groups of employees must be the central goal of data education frameworks.

Moreover, emphasizing performance measures (e.g., knowledge and skills in using different tools) while overlooking perceptual and individual factors (e.g., confidence and trust) hinders data literacy programs from achieving their outcomes [22]. The ultimate goal of these programs should be to help employees build self-efficacy with data while increasing their positive perception of data use. This demands a paradigm shift in data literacy programs by encouraging learners to learn independently and

develop their confidence with data. This includes addressing potential misconceptions to encourage end-users to understand the value of data literacy and make a commitment to using it in their work. When employees pinpoint the value of data and its application themselves, they are more open to learn and participate in future data literacy initiatives.

## 3 A Model for Data Literacy Education

At the core of data literacy is the human ability to learn; literacy involves a person trying to learn something new. To create an education framework for increasing data literacy, TLT and ELT play a complementary role in addressing the challenges to data education for end-users discussed above.

Transformative learning refers to the process of effecting change in a *frame of reference*. Experiential learning defines "learning as the process whereby knowledge is created through the *transformation of experience*. Knowledge results from the combination of grasping and transforming experience" [23, p. 41]. In putting these theories together, the present model combines a behavioral learning approach with a holistic and integrative perspective [23].

### 3.1 Transformative Learning Theory

The TLT argues that we must understand the meaning of our experiences in order to internally validate them [24]. The goal of this theory, as it relates to our research, is to use self-driven critical reasoning to justify the value of data literacy, so that end-users in organizations develop internal motivation to participate in the learning experience. Employees tend to reject ideas that do not fit their preconceptions and assumptions, labeling them as unworthy of consideration or irrelevant. However, when circumstances permit, employees move towards a frame of reference that is more inclusive and integrative of their own experiences [25]. TLT was developed initially to help create such circumstances.

Our study confirms that many employees rely on their experience, intuition, and historical patterns when making decisions. Their unfamiliarity with the value of data in combination with evolving data tools has the potential to discourage them from changing their decision-making routines. As a result, employees use limitations of data and data tools as excuses for not dissecting decision-making problems altogether [7]. Furthermore, employees may ignore the data since they do not trust it at face value, or it does not align with their intuition from their frame of reference [7].

TLT offers four strategies to tackle these problems. These strategies accommodate adult learning needs and habits that directly affect their

frame of reference (e.g., cognition, value, and experience). These strategies are related to learning disorientation, reflection, discourse, and action.

Personal perspectives act as filters that influence our experiences and actions. Disorienting dilemmas can challenge learners to rethink their perspectives and rebuild them through discourse and reflection. TLT claims that learners use experiences as a starting point, meaning that learning is the process of using prior experience to generate a new or revised understanding [25]. Moreover, according to Mezirow [25], learning includes examining assumptions and changing one's beliefs and behaviors through reflection and dialogue. Thus, it is through critical reflection and discourse that transformative learning can take place. A perspective transformation can lead to new a frame of reference that is more inclusive, differentiating, permeable, critically reflective, and integrative of experiences [25]. TLT in this sense helps learners become more autonomous thinkers through the evaluation of their values, meanings, and purpose rather than uncritically acting on potential biases.

Action follows discourse and reflection. Autonomous action (in the case of this study, using data) occurs after the development of critical awareness. For example, after critical reflection and dialogue with their colleagues, employees can become aware of limitations associated with their decision-making habits in the past and then take actions to address those limitations. With this new perspective, employees are more willing to think, learn, and act with data driven insight [4].

TLT transforms learners' collective frames of reference [4]. It requires critical reflection on conventional assumptions, validation of learner understanding through discursive inquiry, and finally action that reflects this transformation. If successful, employees that act on the new perspective may encourage others to do so as well. This approach ultimately envisions educated employees engaged in continuous collaborative inquiry while still learning autonomously themselves [4], shifting the culture of the organization towards one with a greater emphasis on learning and growth overall.

In an organizational context, successful implementation of a learning program depends on the organizational leaders engaging employees in informed, effective, and appropriate learning action. However, TLT lacks a strategy that promotes such engagement, as it is more focused on knowledge acquisition and manipulation of what has been learned. To address this shortcoming, this study proposes the inclusion of both TLT and ELT in data education frameworks.

## 3.2 Experiential Learning Theory

ELT is a method of effective learning where the student learns by experiences. Individuals retain information better through experiential learning because they have the opportunity to apply their knowledge in a learning environment. A common model used to implement ELT is Dewey's model of learning [26]. This model involves: (a) observation of surrounding conditions, (b) knowledge of what has happened in similar situations in the past; and (c) judgment. The purpose of this is give foresight to users so that observation and judgement of the situation are performed, rather than immediate action [6]. This model places an emphasis on refraining from using experiences alone as a deciding factor, but also knowledge of concepts, previous observations and actions. Here, the postponing of immediate action by impulse is critical for the judgement and observation to take place, in addition to the collective factors above giving direction to impulse. The integration of these "opposing but symbiotically related processes" allows for informed decisions [6], [27]

ELT addresses the limitations of TLT by offering strategies to construct knowledge and meaning from real life experience. In the context of our work, ELT can offer a framework to train employees in natural settings [6]. Contextual observations and judgements play a critical role in driving action as they are the foundation of reflection suggested by TLT [23].

## 4 Case Study

In order to gain insight into the dynamics of data literacy and implementation within an organizational setting, our team conducted a series of semi-structured interviews with the team leads of two units working with data and leading data literacy initiatives within the U.S. Department of Defense. For the purpose of discretion, the specific teams will not be specified when stating our findings. They will be referred to as Team X and Team Y. Fourteen in-depth interviews were conducted online via Zoom due to limitations related to the COVID-19 pandemic. These interviews aimed to understand the scope of the teams' training models and processes for increasing data literacy as well as explore the employees' reactions to those initiatives.

Following the interviews, we conducted a thematic coding to identify the successes and challenges associated with each component of their data literacy programs. The categorization of these factors and realignment of them with TLT and ELT key strategies allowed us to develop four key contextual strategies informing future data literacy training programs.

Both teams defined data literacy generally as having the knowledge, skills, and motivation to leverage data to make better decisions. Building on this, the interviewees emphasized the need for employees to be able to clean, read, and extract data comfortably. The ability to share data with others within an organization provides greater insight for all those involved. Interviews revealed a divide in leadership in terms of appropriate measures of data literacy in end-user employees. While leaders were more outcome-oriented (e.g., saving resources), middle-range managers and employees focused more on the procedural values of data literacy (e.g., completing tasks such as reporting and forecasting).

The interviews also revealed that, when making decisions, many employees and managers trust their intuition over the information provided by data. Those in higher positions, sometimes outside of data-driven roles, are outcome focused for the purpose of filling agendas specific to their role. However, when provided with data, a lack of trust arises if results do not match their intuition. One team stated that organizational leaders expect employees to make data driven decisions automatically; however, the processes and resources for getting employees ready to make these of decisions do not properly support them in doing so.

Interviewees indicated a divide in the expectations from end-users. Team X suggested that data literacy is a collective organizational quality. Interviewees on this team explained that end-users could be data literate individually through knowing definitions related to data use, why it is important, where it came from, and how to implement it. However, they might not collectively be capable to make data-driven decisions. Members of Team Y offered another view, sharing that leaders should be responsible to provide end-users information on how to leverage and interpret data, while end-users should be responsible for adhering and applying the information provided by their superiors.

The interviews also revealed that end-users might show little interest in data (and data literacy) if they would not experience its immediate value firsthand. To improve this, employees must learn why data improve insight and decision making. This would motivate end-users to use and eventually develop trust in data. One of the interviewees identified that, when end-users fail to trust or utilize data, it is the responsibility of data literacy programs to find an explanation for their action/inaction, rather than functioning based on assumptions.

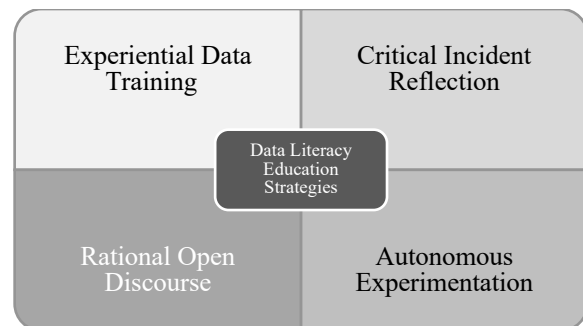
When looking for competencies, positive attitude toward data over mere proficiency in tools was preferred by most interviewees. Members of both teams sought end-users that are willing to commit to

develop data literacy, reflecting that teaching end-users how to use tools is far less complicated than convincing end-users of their necessity. Interviewees also noted the need to motivate users to learn how to use different tools more independently due to the rapid evolution of data collection and tools.

While seen as less important than a positive attitude toward learning data literacy, knowledge of data tools was also emphasized. Common data tools between the two teams include: *Tableau* (data access, sharing, and analysis), *Python* (analysis), *Jira*, *Confluence*, and *SharePoint* (collaboration). However, the results of the interview suggest that proficiency in using data tools does not mean data literacy per se. Interviewees explained that data literacy includes an understanding of the value of data in decision making processes, not just the ability to follow data analysis procedures. In other words, one can be proficient in data literacy but still not have a solid understanding of the purpose of the data. Therefore, interviewees showed a preference to hire data savvy and literate employees motivated to use and trust data to make data driven decisions, not those with competencies in tools.

## 5 Discussion

We summarized our findings in a model comprising four strategies: *experiential data training*, *critical incident reflection*, *rational open discourse*, and *autonomous experimentation* (Figure 1).



**Figure 1. Data literacy education strategies**

Experiential data training encourages the use of one's own experience to validate and learn both the value and applications of data. Critical incident reflection involves examining assumptions and challenging initiative decision making [28]. Rational open discourse promotes a systematic and ongoing dialogue between key stakeholders about data applications. Lastly, autonomous experimentation empowers employees to experiment with data tools and their applications in natural settings before formal adoption.

Together, these strategies offer a systematic approach to data literacy education that can transform and support organizations to better recruit, train, motivate, and engage data savvy employees.

## 5.1 Experiential Data Training

In a traditional data education training environment, end-users memorize procedures and the steps to preform them. Deviating from this model, experiential data training encourages individuals to develop a deeper understanding of data programs and applications, teaching end-users how to use programs and applications to solve new problems. This framework for data education is informed by experiential learning, which links education, work, and personal development. Joining these different sectors of life together allows end-users to feel more confident when learning and experimenting with new data techniques. Some interviewees noted that many end-users tend to trust their intuition more than data. This causes end-users to make faulty decisions without proper data supporting their reasoning. Experiential data training thus addresses the lack of trust many end-users face when it comes to data. Reflections from different interviewees support this argument, explaining that organizational leaders should prioritize teaching the process over successful outcomes. This encourages end-users to critically think when utilizing data. Being exclusively outcome-focused hinders the opportunity for data users to properly understand data in context and learn about their own assumptions and biases related to data. More specifically, user focus on successful outcomes prevents end-users from seeing the need for data when intuition-driven decisions generate equally successful results. As a result, users fail to see the need to increase their own data literacy.

Experiential data training is a key facet of data literacy education as it allows end-users to continuously learn and engage with data and data tools over time as part of day-to-day routines. End-users can learn more effectively about data and data tools when they can make mistakes while experiencing possibilities and limitations at the same time. Instead of dry, formulaic formal trainings, end-users are encouraged to explore data tools on their own, allowing continuous learning critical to constantly shifting data practices and tools. An organization solely focusing on outcomes and results may hinder an end-user's learning instead of promoting lifelong experiential learning. Therefore, we expect

***Proposition 1: Experiential Data Training can improve Data Literacy in organizational settings.***

## 5.2 Critical Incident Reflection

Critical reflection normalizes challenging the validity of intuition not backed by data. Becoming critically self-aware of presuppositions involves challenging established and habitual patterns of expectation, those perspectives one uses to make sense out of encounters with the world, others, and oneself.

Informed by TLT and ELT, we propose the implementation of a system that encourages end-users (both employees and management) to reflect on their own experiences with data and data tools. TLT suggests that challenging employees' understandings of data helps them revisit their assumptions and build a new frame of reference. This also helps organizations to gauge the end-users' data literacy and their progress in a more meaningful way than merely evaluating outcomes. However, the reflection process should be systematic in order to improve organizational data literacy as a whole. For example, end-users should be asked to recall specific instances of data applications and their reactions to them. The reflections should be shared with management in a systematic way as well. When reflections are encouraged and feedback is embraced, end-users can develop confidence and trust in their abilities to appraise data and their applications in different contexts. As end-users progress, they can react more rationally to what they face in new settings.

Both ELT and TLT encourage learners to reflect on past experiences through the process of reflective journaling. When adopted by data literacy education initiatives, reflective journaling can help end-users to practice and present a collection of incidents that can be used to improve data use. This leads to normalizing the use and value of data as it grows within an organization. This documentation can include what occurred, what actions were taken, how the experience enhanced learning, and what the individual took away from the learning experience. This approach to reflection provides a structure for speculation and metacognition, as well as synthesis of events experienced when utilizing data to produce solutions. This can be beneficial to trainers in charge of increasing data literacy as it can provide explanations for unsuccessful training practices, leading to more robust training programs. Therefore, critical incident reflection allows users to make observations and judgements to make informed decisions. Thus,

***Proposition 2: Critical Incident Reflection can improve Data Literacy in organizational settings.***

### 5.3 Rational Open Discourse

TLT suggests individual reflection without group discourse cannot be fully effective. Hence, in order to encourage data literacy organization wide, data literacy needs, challenges, and opportunities must be communicated in a way that can be consumed amongst all members of an organization, not just data analysts and scientists. These discussions can trigger critical thinking and foster open dialogue between data stakeholders. We propose creating a safe environment where even the most logical data solutions can be questioned or challenged. In order to engage end-users in reflective discourse, alternative perspectives must be examined to promote dialectical thinking—a metacognitive trait. After the presentation of multiple perspectives, end-users should then be encouraged to synthesize the information and arrive at a logical conclusion regardless of their original/intuitive viewpoint. This process requires the discussion and evaluation of results, an openness to alternative approaches, and the comparison of various solutions.

Furthermore, TLT encourages systematic and ongoing dialogue between key stakeholders about data applications. For example, organizations have a variety of tools to choose from, and those tools need to be effective when it comes to data collection and sharing systems. Open discussion about the selection and deployment of data tools ensures that data collection systems are both practical and feasible for all members of a project. After implementation, a dialogue between all stakeholders could also allow more desirable and feasible adjustments. Open discourse, when planned and organized, can pave the way for data literacy programs to communicate the needs, priorities, and challenges of an organization in an inclusive manner, allowing all stakeholders the ability to participate. Therefore, rational open discourse allows a user's own views to be challenged to come to informed actions. Thus,

***Proposition 3: Rational Open Discourse can improve Data Literacy in organizational settings.***

### 5.4 Autonomous Experimentation

Supported by TLT principles and self-determination theory [29], we identified autonomous experimentation with data as a mechanism to improve data literacy among end-users. ELT also suggests autonomous experimentation to internalize learning. Granting end-users the freedom, tools, and time to experiment with data and data tools can expand their knowledge independently. Autonomous experimentation involves the allocation of time for employees to navigate and utilize data applications as

opposed to formal trainings that teach uniform data application techniques. This is because cognitive and metacognitive learning is a highly individual process. Autonomy allows learners to delve into the plethora of information and possibilities on their own time. In allowing end-users to think and explore creatively, data education programs can empower individuals to become comfortable with using and understanding data and data tools faster. This promotes a culture of data-informed innovation and risk-taking that ultimately contributes to collective data literacy.

Autonomous experimentation can be promoted by giving employees the discretion to perform projects defined and initiated independently beyond the scope of their job responsibilities. Time can be allocated for employees to engage in exploratory thinking and experimentation with data tools [30]. This time gives room for data innovation by helping end-users develop deeper understandings of data and data applications. While different users may arrive at different findings, the combined results contribute to more personalized learning experiences and thus greater collective data literacy. Our findings suggest that autonomy through organizational or departmental flexibility allows for experimentation via data tools and with data itself. This gives employees the opportunity to go beyond trainings and work independently, which, in turn, increases their intrinsic motivation and propensity to use data. Therefore, we expect

***Proposition 4: Autonomous Experimentation can improve Data Literacy in organizational settings.***

## 6 Implications

This research suggests the efficacy of experiential data training, critical incident reflection, rational open discourse, and autonomous experimentation in redesigning data literacy education programs. TLT and ELT models informed the selection of these key strategies in the implementation of data literacy education programs and formation of organizational data literacy—or, as we call it, ‘data swagger’. Our proposed model provides a roadmap for organizations to engage, motivate, and train the next generation of data literate end-users better than traditional one-way training models. We contextualized the theories presented to inform both future research and practice. This contextualization suggests that an experience-centered, reflection-driven, and cooperative approach to data literacy has the potential to produce more data savvy users. Future research should thus utilize the proposed strategies to create more specific, contextual, and effective training programs, as well as evaluate each strategy in practice.



We suggest, for effective implementation, emphasis should be placed on end-users' personal experiences with data and data tools. This 'employee-first' approach implies that, in planning data literacy programs, employees' needs should come before procedural requirements or organizational priorities. Assuming the successful implementation of our propositions, employees should be more motivated and engaged with data education programs, exploring the uses and nuances of data beyond the provided training materials.

In order to emphasize the practicality of our framework, we have created different example scenarios in which each of the four suggestions from our model can be implemented by data literacy trainers. Experiential data training can be implemented when organizations have access to different use cases that are understandable and meaningful to most participants in the training. Use cases can be initially planned by the trainers or brought to the training by trainees. Since experiential data training is a process of active learning, trainers should create an environment in which employees are encouraged to ask questions, develop their technical skills, and make mistakes all while learning new ways to interact with data. Trainers should also ask trainees to seek a variety of outcomes when trying to find a solution using data. For example, instead of asking them to find an acceptable solution following the instructions, instructors should task trainees to review a set of use cases and then explain why the organization should adopt one over the others.

Critical incident reflection can increase engagement when incorrect assumptions hinder data literacy programs from achieving their desired outcomes. Trainers should ask participants to document major incidents and experiences, positive or negative, with data, tools, and applications. Afterwards, participants should reflect on the individual, situational, and organizational factors contributing to such outcomes. This exercise encourages trainees to think critically about data literacy and thus motivates them to be more open to opportunities and challenges with data in the future.

Rational open discourse can be coordinated by the trainers after the implementation of critical incident reflection. This discourse should involve trainees sharing their experiences and seeking feedback from peers in an open, transparent, and non-threatening learning environment. This creates a sense of community among the trainees, which enhances cooperative learning in both the short and long term. Trainers can also benefit from participation in the

discourse in three ways: (a) learning about misunderstandings/misconceptions and addressing them in a timely manner, (b) identifying teaching by example opportunities, and (c) adjusting the course of trainings or the resources offered.

Autonomous experimentation can take data literacy programs to a new level. Providing testbeds and a sandbox for trainees can be a practical solution to encourage them to practice and experiment with what they have learned. It is also a good practice to co-experiment with data and data tools in a group setting, looking back on outcomes together. This encourages trainees to use data and data tools innovatively. In addition, organizational leaders can encourage experimentation and reward creative solutions.

Lastly, we recommend a process-based approach to data literacy training over an outcome-based approach focused merely on procedural use. We argue that data literacy is the result of end-users' engagement in a series of well-planned learning activities through which they develop both contextual knowledge and motivation. Therefore, hiring data savvy employees would not be enough in creating a data savvy team, much less to realize the full value from such environment.

## **7 Limitations and Future Research**

This study was limited to a set of interviews with members of two teams within one organization. Therefore, the results are not generalizable to other contexts without further examination. Our framework is, however, grounded in well-established educational theories to be adopted and tested in different types of organizational settings. This study is also limited in scope. For example, we did not consider the content of data literacy programs—their quality and utility. Moreover, this research does not provide a representative overview of all existing problems in recruiting, training, motivating, and engaging data savvy employees. Rather, this is an exploratory study demonstrating the primary barriers to increasing data literacy within an organization. To move past those barriers, we propose a framework to assist future research and practice.

Our data collection was also limited to the members of two units that were responsible for data management and analytics, trainings, and supports. Future research can include more additional sources of data and interviews with both organizational leaders as well as lower tier end-users. It is essential to observe the divide in perspectives and refine the proposed strategies. Furthermore, more empirical research is

needed on the benefits and limitations of each proposed strategy. Additional research is also needed to understand the benefits of rational open discourse between key stakeholders about data applications and the benefits of autonomy and experimentation within a data-driven environment. Future research should experimentally test our propositions to assess the causal relationship between the implementation of these four strategies and improved data literacy amongst end-users. Documenting and comparing different mechanisms to implement these strategies also deserves some attention.

Finally, we believe it is important to reemphasize the emerging understanding of data literacy. Although the concept of data literacy is not new, the contextualization of how people and organizations can become data literate is a topic that is not widely understood. The nature of our research leaves much to be expanded upon, so future researchers may develop more detailed implementation strategies by using our framework as a frame of reference.

## 8 References

- [1] K. Abhari, A. Vomero, and E. Davidson, "Psychology of Business Intelligence Tools: Needs-Affordances-Features Perspective," 2020, doi: 10.24251/hicss.2020.602.
- [2] J. Don-Carolis, "Business Insights: Global," *TrueCue*, 2021. .
- [3] G. Gimpel, "Bringing dark data into the light: Illuminating existing IoT data lost within your organization," *Bus. Horiz.*, vol. 63, no. 4, pp. 519–530, Jul. 2020.
- [4] J. Mezirow, "Transformative learning and social action: A response to inglis," *Adult Educ. Q.*, vol. 49, no. 1, pp. 70–72, Sep. 1998, doi: 10.1177/074171369804900109.
- [5] A. Kolb and D. Kolb, "Experiential Learning Theory as a Guide for Experiential Educators in Higher Education," *Exp. Learn. Teach. High. Educ.*, vol. 1, no. 1, Jun. 2017.
- [6] D. A. Kolb, *Experiential learning: experience as the source of learning and development*. Prentice-Hall, 1984.
- [7] J. Mezirow, "Transformative Learning as Discourse," *J. Transform. Educ.*, vol. 1, no. 1, pp. 58–63, Jul. 2003.
- [8] A. N. Markham, "Taking Data Literacy to the Streets: Critical Pedagogy in the Public Sphere," *Qual. Inq.*, vol. 26, no. 2, pp. 227–237, Feb. 2020.
- [9] J. Cook, D. Bedford, and S. Mandia, "Raising climate literacy through addressing misinformation: Case studies in agnotology-based learning," *Journal of Geoscience Education*, vol. 62, no. 4. National Association of Geoscience Teachers Inc., pp. 296–306, Nov. 2014..
- [10] *Data Science for Undergraduates*. National Academies Press, 2018.
- [11] E. Eryurek, U. Gilad, V. Lakshmanan, A. Kibunguchy, and J. Ashdown, *Data Governance: The Definitive Guide*. O'Reilly Media, 2021.
- [12] J. Watts, L. Sare, and D. E. Hubbard, "Collaborative Data Literacy Education for Research Labs: A Case Study at a Large Research University," 2021.
- [13] P. Attewell and D. Monaghan, "1. What Is Data Mining?," *Data Min. Soc. Sci.*, no. 2004, pp. 3–12, 2019.
- [14] Y. Yin, I. Gurvich, S. McReynolds, D. Seys, and J. A. Van Mieghem, "Learning by doing versus learning by viewing: An empirical study of data analyst productivity on a collaborative platform at eBay," *Proc. ACM Human-Computer Interact.*, vol. 2, no. CSCW, 2018.
- [15] P. Mikalef, M. N. Giannakos, I. O. Pappas, and J. Krogstie, "The human side of big data: Understanding the skills of the data scientist in education and industry," *IEEE Glob. Eng. Educ. Conf. EDUCON*, vol. 2018-April, pp. 503–512, 2018.
- [16] J. Windt, D. Taylor, D. Little, and B. C. Sporer, "Making everyone's job easier. How do data scientists fit as a critical member of integrated support teams?," *Br. J. Sports Med.*, vol. 55, no. 2, pp. 73–75, 2021, doi: 10.1136/bjsports-2020-102938.
- [17] B. P. Alant and O. O. Bakare, "A case study of the relationship between smallholder farmers' ICT literacy levels and demographic data w.r.t. their use and adoption of ICT for weather forecasting," *Heliyon*, vol. 7, no. 3, Mar. 2021.
- [18] K. Vassakis, E. Petrakis, and I. Kopanakis, "Big Data Analytics: Applications, Prospects and Challenges," *Lect. Notes Data Eng. Commun. Technol.*, vol. 10, 2018.
- [19] W. A. Günther, M. H. Rezazade Mehrizi, M. Huysman, and F. Feldberg, "Debating big data: A literature review on realizing value from big data," *J. Strateg. Inf. Syst.*, vol. 26, no. 3, pp. 191–209, Sep. 2017.
- [20] S. Gupta, A. K. Kar, A. Baabdullah, and W. A. A. Al-Khowaiter, "Big data with cognitive computing: A review for the future," *Int. J. Inf. Manage.*, vol. 42, pp. 78–89, Oct. 2018.
- [21] G. Dishon, "New data, old tensions: Big data, personalized learning, and the challenges of progressive education," *Theory Res. Educ.*, vol. 15, no. 3, pp. 272–289, Nov. 2017.
- [22] V. Dhar and A. Sundararajan, "Information technologies in business: A blueprint for education and research," *Inf. Syst. Res.*, 2007, doi: 10.1287/isre.1070.0126.
- [23] D. A. Kolb, *Experiential Learning: Experience as The Source of Learning and Development*. 1984.
- [24] M. C. Clark and A. L. Wilson, "Context and rationality in mezirow's theory of transformational learning," *Adult Educ. Q.*, vol. 41, no. 2, pp. 75–91, Sep. 1991.
- [25] J. M. Dirkx, J. Mezirow, and P. Cranton, "Musings and Reflections on the Meaning, Context, and Process of Transformative Learning: A Dialogue Between John M. Dirkx and Jack Mezirow," *J. Transform. Educ.*, vol. 4, no. 2, pp. 123–139, Jul. 2006.
- [26] J. Dewey, *Experience and Nature*. Dover Publications, 1958.
- [27] J. Dewey, "Experience and education," *Educ. Forum*, vol. 50, no. 3, pp. 242–252, Sep. 1986.
- [28] D. L. Gilstrap and J. Dupree, "A Regression Model of Predictor Variables on Critical Reflection in the Classroom: Integration of the Critical Incident Questionnaire and the Framework for Reflective Thinking," *J. Acad. Librariansh.*, vol. 34, no. 6, pp. 469–481, Nov. 2008, doi: 10.1016/j.acalib.2008.09.019.
- [29] E. L. Deci, A. H. Olafsen, and R. M. Ryan, "Self-Determination Theory in Work Organizations: The State of a Science," *Ssrn*, vol. 4, no. April, pp. 19–43, 2017.
- [30] A. Burcharth, M. Præst Knudsen, and H. A. Søndergaard, "The role of employee autonomy for open innovation performance," *Bus. Process Manag. J.*, vol. 23, no. 6, pp. 1245–1269, 2017.